


[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)
**Search:**  The ACM Digital Library  The Guide

## THE ACM DIGITAL LIBRARY

[Feedback](#)

((document clustering entropy probability))  
 Published before December 2003  
 Terms used:  
**document clustering entropy probability**

Sort  
results  
by

Display  
results

Save these results to a Binder [Save](#) [Refine](#)  
 Open results in a new window [Open](#) [Try this search in The ACM Guide](#)

---

Results 1 - 50 of 105 Result page: 1 [2](#) [3](#) [next](#)

[>>](#)

### 1 [A method of cluster-based indexing of textual data](#)

Akiko Aizawa

August 2002 [Proceedings of the 19th international conference on Computational linguistics - Volume 1, Volume 1](#)

### 2 [MPEG-4: an object-based multimedia coding standard supporting mobile applications](#)

Atul Puri, Alexandros Eleftheriadis

June 1998 [Mobile Networks and Applications, Volume 3 Issue 1](#)

### 3 [A neural probabilistic language model](#)

Yoshua Bengio, Réjean Ducharme, Pascal Vincent, Christian Janvin  
 March 2003 [The Journal of Machine Learning Research, Volume 3](#)

### 4 [Latent dirichlet allocation](#)

David M. Blei, Andrew Y. Ng, Michael I. Jordan  
 March 2003 [The Journal of Machine Learning Research, Volume 3](#)

**5 Simple and accurate feature selection for hierarchical categorisation**



Wahyu Wibowo, Hugh E. Williams

November DocEng '02: Proceedings of the 2002 ACM symposium on Document  
2002 engineering

**6 Distribution-based pruning of backoff language models**

Jianfeng Gao, Kai-Fu Lee

October ACL '00: Proceedings of the 38th Annual Meeting on Association for  
2000 Computational Linguistics

**7 Data mining for hypertext: a tutorial survey**



Soumen Chakrabarti

January ACM SIGKDD Explorations Newsletter, Volume 1 Issue  
2000 2

**8 Cross-lingual C\*ST\*RD: English access to Hindi information**



Anton Leuski, Chin-Yew Lin, Liang Zhou, Ulrich Germann, Franz Josef Och, Eduard Hovy  
September ACM Transactions on Asian Language Information Processing  
2003 (TALIP), Volume 2 Issue 3

**9 Enhanced topic distillation using text, markup tags, and hyperlinks**



Soumen Chakrabarti, Mukul Joshi, Vivek Tawde

September SIGIR '01: Proceedings of the 24th annual international ACM SIGIR  
2001 conference on Research and development in information retrieval

**10 Mining the web for answers to natural language questions**



Dragomir R. Radev, Hong Qi, Zhiping Zheng, Sasha Blair-Goldensohn, Zhu Zhang,  
Weiguo Fan, John Prager

October CIKM '01: Proceedings of the tenth international conference on Information  
2001 and knowledge management

**11 Compression of scan-digitized Indian language printed text: a soft pattern  
matching technique**



U. Garain, S. Debnath, A. Mandal, B. B. Chaudhuri

November DocEng '03: Proceedings of the 2003 ACM symposium on Document  
2003 engineering

**12 Entropy-based link analysis for mining web informative structures**



Hung-Yu Kao, Ming-Syan Chen, Shian-Hua Lin, Jan-Ming Ho

November CIKM '02: Proceedings of the eleventh international conference on  
2002 Information and knowledge management

**13 An extensive empirical study of feature selection metrics for text  
classification**

George Forman

March The Journal of Machine Learning Research, Volume 3  
2003

**14 Activity detection for information access to oral communication**

Klaus Ries, Alex Waibel

March HLT '01: Proceedings of the first international conference on Human  
2001 language technology research

**15** A divisive information theoretic feature clustering algorithm for text classification

Inderjit S. Dhillon, Subramanyam Mallela, Rahul Kumar  
March The Journal of Machine Learning Research, Volume 3  
2003

**16** Topic-based document segmentation with probabilistic latent semantic analysis

 Thorsten Brants, Francine Chen, Ioannis Tsachantaridis  
November CIKM '02: Proceedings of the eleventh international conference on  
2002 Information and knowledge management

**17** ReCoM: reinforcement clustering of multi-type interrelated data objects

 Jidong Wang, Huajun Zeng, Zheng Chen, Hongjun Lu, Li Tao, Wei-Ying Ma  
July SIGIR '03: Proceedings of the 26th annual international ACM SIGIR conference  
2003 on Research and development in informaion retrieval

**18** Learning with mixtures of trees

Marina Meila, Michael I. Jordan  
September The Journal of Machine Learning Research, Volume  
2001 1

**19** Progress in Picture Processing: 1969--71

 Azriel Rosenfeld  
June ACM Computing Surveys (CSUR), Volume 5 Issue  
1973 2

**20** Bit-sliced index arithmetic

 Denis Rinfret, Patrick O'Neil, Elizabeth O'Neil  
June ACM SIGMOD Record, Volume 30 Issue  
2001 2

**21** Clustering documents in a web directory

 Giordano Adamo, Paolo Avesani, Diego Sona  
November WIDM '03: Proceedings of the 5th ACM international workshop on Web  
2003 information and data management

**22** Trace-driven system-level power evaluation of system-on-a-chip peripheral cores

 Tony D. Givargis, Frank Vahid, Jörg Henkel  
January ASP-DAC '01: Proceedings of the 2001 conference on Asia South Pacific  
2001 design automation

**23** Generalized algorithms for constructing statistical language models

Cyril Allauzen, Mehryar Mohri, Brian Roark  
July ACL '03: Proceedings of the 41st Annual Meeting on Association for  
2003 Computational Linguistics - Volume 1, Volume 1

**24** Inducing German semantic verb classes from purely syntactic subcategorisation information

Sabine Schulte im Walde, Chris Brew  
July ACL '02: Proceedings of the 40th Annual Meeting on Association for  
2001 Computational Linguistics

**25 Preserving peer replicas by rate-limited sampled voting**

 Petros Maniatis, David S. H. Rosenthal, Mema Roussopoulos, Mary Baker, TJ Giuli, Yanto Muliadi  
October 2003 SOSP '03: Proceedings of the nineteenth ACM symposium on Operating systems principles

**26 Data-driven evolution of data mining algorithms**

 Padhraic Smyth, Daryl Pregibon, Christos Faloutsos  
August 2002 Communications of the ACM, Volume 45 Issue 8

**27 Integrating the document object model with hyperlinks for enhanced topic distillation and information extraction**

 Soumen Chakrabarti  
April 2001 WWW '01: Proceedings of the 10th international conference on World Wide Web

**28 Document clustering based on non-negative matrix factorization**

 Wei Xu, Xin Liu, Yihong Gong  
July 2003 SIGIR '03: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval

**29 Clinical and financial outcomes analysis with existing hospital patient records**

 R. Bharat Rao, Sathyakama Sandilya, Radu Stefan Niculescu, Colin Germond, Harsha Rao  
August 2003 KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining

**30 Mining the peanut gallery: opinion extraction and semantic classification of product reviews**

 Kushal Dave, Steve Lawrence, David M. Pennock  
May 2003 WWW '03: Proceedings of the 12th international conference on World Wide Web

**31 Client-server computing in mobile environments**

 Jin Jing, Abdelsalam Sumi Helal, Ahmed Elmagarmid  
June 1999 ACM Computing Surveys (CSUR), Volume 31 Issue 2

**32 Probabilistic text structuring: experiments with sentence ordering**

Mirella Lapata  
July 2003 ACL '03: Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 1, Volume 1

**33 Frequency estimates for statistical word similarity measures**

Egidio Terra, C. L. A. Clarke  
May 2003 NAACL '03: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - Volume 1, Volume 1

**34 Polyphonic music modeling with random fields**

 Victor Lavrenko, Jeremy Pickens  
November 2003 MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia

**35 Multi-resolution disambiguation of term occurrences**  
 Einat Amitay, Rani Nelken, Wayne Niblack, Ron Sivan, Aya Soffer  
November 2003 CIKM '03: Proceedings of the twelfth international conference on Information and knowledge management

**36 Getting more mileage from web text sources for conversational speech language modeling using class-dependent mixtures**  
Ivan Bulyko, Mari Ostendorf, Andreas Stolcke  
May 2003 NAACL '03: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology: companion volume of the Proceedings of HLT-NAACL 2003--short papers - Volume 2, Volume 2

**37 Summarizing scientific articles: experiments with relevance and rhetorical status**  
Simone Teufel, Marc Moens  
December 2002 Computational Linguistics, Volume 28 Issue 4

**38 Clustering verbs semantically according to their alternation behaviour**  
Sabine Schulte im Walde  
July 2000 Proceedings of the 18th conference on Computational linguistics - Volume 2, Volume 2

**39 VideoQA: question answering on news video**  
 Hui Yang, Lekha Chaisorn, Yunlong Zhao, Shi-Yong Neo, Tat-Seng Chua  
November 2003 MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia

**40 Document clustering with cluster refinement and model selection capabilities**  
 Xin Liu, Yihong Gong, Wei Xu, Shenghuo Zhu  
August 2002 SIGIR '02: Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval

**41 Toward a unified approach to statistical language modeling for Chinese**  
 Jianfeng Gao, Joshua Goodman, Mingjing Li, Kai-Fu Lee  
March 2002 ACM Transactions on Asian Language Information Processing (TALIP), Volume 1 Issue 1

**42 Using web structure for classifying and describing web pages**  
 Eric J. Glover, Kostas Tsoutsouliklis, Steve Lawrence, David M. Pennock, Gary W. Flake  
May 2002 WWW '02: Proceedings of the 11th international conference on World Wide Web

**43 Statistical acquisition of content selection rules for natural language generation**  
Pablo A. Duboue, Kathleen R. McKeown  
July 2003 Proceedings of the 2003 conference on Empirical methods in natural language processing - Volume 10, Volume 10

**44 Semantic language models for topic detection and tracking**

Ramesh Nallapati

May NAACL '03: Proceedings of the 2003 Conference of the North American  
2003 Chapter of the Association for Computational Linguistics on Human  
Language Technology: Proceedings of the HLT-NAACL 2003 student  
research workshop - Volume 3, Volume 3

**45 Instruction-based system-level power evaluation of system-on-a-chip peripheral  
cores**

Tony D. Givargis, Frank Vahid, Jörg Henkel

September ISSS '00: Proceedings of the 13th international symposium on System  
2000 synthesis

**46 Sufficient dimensionality reduction**

Amir Globerson, Naftali Tishby

March The Journal of Machine Learning Research, Volume  
2003 3

**47 Multi-level background initialization using Hidden Markov Models**

Marco Cristani, Manuele Bicego, Vittorio Murino

November IWVS '03: First ACM SIGMM international workshop on Video  
2003 surveillance

**48 Information-theoretic co-clustering**

Inderjit S. Dhillon, Subramanyam Mallela, Dharmendra S. Modha

August KDD '03: Proceedings of the ninth ACM SIGKDD international conference on  
2003 Knowledge discovery and data mining

**49 Automatic verb classification based on statistical distributions of argument  
structure**

Paola Merlo, Suzanne Stevenson

September Computational Linguistics, Volume 27 Issue 3  
2001

**50 Towards NIC-based intrusion detection**

M. Otey, S. Parthasarathy, A. Ghoting, G. Li, S. Narravula, D. Panda

August KDD '03: Proceedings of the ninth ACM SIGKDD international conference on  
2003 Knowledge discovery and data mining

---

Results 1 - 50 of 105 Result page: 1 [2](#) [3](#) [next](#)

>>

The ACM Portal is published by the

Association for Computing Machinery. Copyright © 2008 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  Adobe Acrobat

 QuickTime

 Windows Media Player

 Real Player